Artificial Intelligence in Diagnosis of Oral Diseases: A Systematic Review

Shaul Hameed Kolarkodi¹, Khalid Zabin Alotaibi²

ABSTRACT

Aim: To understand the role of Artificial intelligence (AI) in oral radiology and its applications.

Background: Over the last two decades, the field of AI has undergone phenomenal progression and expansion. Artificial intelligence applications have taken up new roles in dentistry like digitized data acquisition and machine learning and diagnostic applications.

Materials and Methods: All research papers outlining the population, intervention, control, and outcomes (PICO) questions were searched for in PubMed, ERIC, Embase, CINAHL, database from the last 10 years on first January 2023. Two authors independently reviewed the titles and abstracts of the selected studies, and any discrepancy between the two review authors was handled by a third reviewer. Two independent investigators evaluated all the included studies for the quality assessment using the modified tool for the quality assessment of diagnostic accuracy studies (QUADAS-2).

Review results: After the removal of duplicates and screening of titles and abstracts, 18 full texts were agreed upon for further evaluation, of which 14 that met the inclusion criteria were included in this review. The application of artificial intelligence models has primarily been reported on osteoporosis diagnosis, classification/segmentation of maxillofacial cysts and/or tumors, and alveolar bone resorption. Overall study quality was deemed to be high for two (14%) studies, moderate for six (43%) studies, and low for another six (43%) studies.

Conclusion: The use of AI for patient diagnosis and clinical decision-making can be accomplished with relative ease, and the technology should be regarded as a reliable modality for potential future applications in oral diagnosis.

Keywords: Artificial intelligence, Computer-assisted imaging, Oral diagnostic imaging.

The Journal of Contemporary Dental Practice (2023): 10.5005/jp-journals-10024-3465

INTRODUCTION

Artificial intelligence (AI), also known as machine learning, aims to give machines the ability to simulate intelligent human behavior, once reserved for humans.¹ These advances in AI have a significant impact on dentistry including various subspecialties like orthodontics, oral surgery, implant dentistry, and oral radiology in diagnosis and treatment planning with various advanced software’s. Few recent developments in oral radiology in diagnosis include automated image-based disease detection with other diagnosis-support systems.² The advent of digital radiographs has greatly aided the rapid development of AI in the medical and dental fields, radiographic images can be easily digitally coded and translated into computational language.³ Digital intraoral radiographs, panoramic radiographs, cephalograms, and other extraoral radiographs, computed tomography (CT) scans, and cone-beam computed tomography (CBCT) are an invaluable asset in the development of AI.⁴ Radiologists visually analyze and interpret the findings, AI enables automatic recognition data of complex imaging patterns, along with quantitative analysis.⁵ Therefore, AI has the potential tool in assisting continued research and development, and personalized dental treatment planning based on clinical data analysis thereby achieving more predictable treatment outcomes.⁶

Hence the application of AI in oral radiology plays a major role from routine diagnosis to advanced imaging. It is unclear, how the current literature, AI can aid in the diagnosis, planning, or management of maxillofacial diseases. Therefore, a systematic review was conducted on studies that discussed various modalities of AI, and its application in the detection and diagnosis of oral diseases.

MATERIALS AND METHODS

This present review was completed as per the guidelines by the preferred reporting items for systematic reviews and meta-analyses (PRISMA). No institutional review board permission was required because of the nature of the current investigation.

Focused PICO Question

The population, intervention, control, and outcomes (PICO) question was developed to identify the appropriate studies to answer the following question: “What are the current clinical applications of AI in dental maxillofacial radiology, and how effective is it as a diagnostic tool?” (1) Population: Patient’s clinical images of a dental and maxillofacial region; (2) Intervention: AI
algorithm-based diagnostic model; (3) Comparison: Reference standard, including things like expert osteotomy, implant primary stability, and implant osteotomy; (4) Conclusion: The performance of diagnostic efficiency of the proposed AI model, including accuracy, sensitivity, and specificity.

**Search Strategy**
All research papers outlining the PICO questions were searched for in PubMed, Scopus, CINAHL, and Ebsco databases from the last 10 years on first January 2023. All the pertinent papers were found using the search method. Additionally, all pertinent articles’ reference lists were hand-searched as well. In order to find additional related studies, a manual search was carried out on the hosting publishers (Wiley, ScienceDirect, and Springer) as well as separately on the renowned implant journals.

**Eligibility Criteria**
In order for studies to be considered for inclusion in the systematic review, they must satisfy the inclusion criteria listed below, Articles reported in the English language, studies reporting on radiology- or clinical studies with imaging utilizing AI models in automatic detection of disease, abnormalities or pathologies, and identification of anatomical structures in the dental and maxillofacial region. The exclusion criteria encompassed: Case reports/series, letters to editors, review articles, and inaccessible full-text articles.

**Study Selection**
The research title, abstract, and keywords of the pertinent publications were autonomously reviewed by two investigators (SHK and KZA) to determine their eligibility. Then, all possibly eligible papers’ full texts were retrieved and carefully reviewed to find research that matched all inclusion requirements. A list of the articles to be included in this evaluation was established after any disagreements were discussed with the third reviewer.

**Data Extraction**
Titles and abstracts of the chosen studies were independently evaluated by two authors who screened the titles and select the abstracts for full-text inclusion. Using the Mesh terms, following the inclusion and exclusion criteria all relevant full-text articles were retrieved. Any disagreement between the two review authors was resolved by the third author. The following categories of information were extracted, Author, year of publication, imaging modality, AI-based application technique/model, testing and training data, valid method, and reference standard type.

**Quality Assessment of Individual Studies**
Two independent investigators (SHK and KZA) evaluated all the included studies for the quality assessment using the modified tool for the quality assessment of diagnostic accuracy studies (QUADAS-2). Quality assessment of diagnostic accuracy studies appraises the diagnostic accuracy studies in four domains, Patient selection, index test, reference standard, and flow and timing. Each domain is evaluated in terms of the risk of bias (RoB) and the applicability of the study results. A study was considered as high quality when it had low RoB in at least six out of the seven subdomains, and as low quality when it presented a high or unclear risk in four or more subdomains; the quality of all other studies was deemed as moderate. Any discrepancies were sorted out with the help of a third reviewer.

**Review Results**
The initial search resulted in a total of 187 hits (PubMed: 55, Scopus: 99, CINAHL: 12, and Ebsco: 20). Following the removal of the duplicates and an initial examination of the titles and abstracts, it was decided that 18 of the full texts should be subjected to further evaluation. Of these 18 full texts, only 14 satisfied the inclusion criteria and were therefore considered for this review. The PRISMA flowchart explaining the study selection is presented (Flowchart 1).
Figure 1 shows the clinical applications of AI proposed for the diagnosis of various oral diseases in the reviewed studies. In these studies, four made use of periapical radiographs, four made use of panoramic radiographs, and one made use of both intraoral and panoramic radiographs. Additionally, three made use of CBCT images, and one made use of undescibed dental X-ray images. In addition, one study evaluated the performance of computer-aided diagnostic software by analyzing three images obtained from CT and MRI scanners (Table 1).

In terms of the applications of these AI models, it was found that three studies reported on the diagnosis of osteoporosis, four studies reported on the classification of maxillofacial cysts and/or tumors, three reported on the diagnosis of maxillary sinusitis, and one study was found for the diagnosis of maxillary sinusitis, and one study reported on the most important aspects of the research are outlined (Table 1), which can be found here.

In terms of the procedures that were used to validate the effectiveness of the AI models, five studies made use of split-sample validation, three made use of leave-one-out cross-validation (LOOCV), one made use of 3-fold cross-validation (CV), one used 5-fold CV, two made use of 10-fold CV, two made use of independent sample validation, two made use of multiple validation techniques, and one did not describe (Table 1).

Risk of Bias (RoB) within Studies Using QUADAS-2 Grading

Figures 2 and 3 illustrate the QUADAS-2 quality assessment of RoB and concerns regarding the applicability of the reviewed studies respectively. Overall study quality was deemed to be high for two (14%) studies, moderate for six (43%) studies, and low for another six (43%) studies. Patient selection was rated to exhibit high RoB in 43% of the studies and low RoB in 57% of the studies. All the studies (100%) were judged to have low RoB in terms of index test and flow and timing. Further, the reference standard was found to possess low RoB in 50% of the studies, unclear in 14% of studies, and high RoB in 36% of studies (Fig. 2). Similarly, under concerns regarding the applicability of the studies, 71, 64, and 14% of the studies had high risk in the patient selection, index test, and reference standard subdomains respectively. Low risk of applicability was observed in 29% of studies for patient selection, 36% of studies for index test and 50% of studies for the reference standards. In terms of the applicability of reference standard, 36% of studies showed unclear risk (Fig. 3).

Discussion

The current investigation was carried out in locating specific areas in dental radiology utilizing AI methodologies and tools. In the beginning, images that were used to construct computer-aided programs to assist with the clinical diagnosis were two-dimensional, including periapical and panoramic radiographs. An AI model for the segmentation of maxillofacial cysts was proposed by Abdolali et al. in the year 2016, the model would use CBCT images obtained from patients. After that, many studies tried to develop CBCT-based AI models in order to solve a clinical problem.

The current review shows that there has been an increase in reports on AI techniques for various aspects of oral and maxillofacial imaging, with the majority of studies concentrating on four main applications. These applications include the screening and diagnosis for osteoporosis, screening and classification involving maxillofacial cysts and/or tumors, and detection of periapical diseases (Table 1; Fig. 1).

For osteoporosis and low bone mineral density (BMD) measurements, the role of the radiologist is very much vital in these screening and planning, difficulties faced by the radiologist can be overcome by the applications of AI in the field of oral medicine, including implant dentistry. Patients who have osteoporosis have a higher risk of having bone loss in the marginal area around the implants, and patients who are treated with antiresorptive drugs have a higher risk of developing osteonecrosis in the jaw bones after oral surgical procedures. This review included three studies that reported models of AI to screen and classify normal and osteoporotic subjects using panoramic radiographs, low skeletal BMD, identifying high bone turnover rate and an increased risk for osteoporotic fracture were used as the basis for the models based on narrowing and erosion of the mandibular cortex. The studies were conducted in China, Japan, and the United States. Utilization of AI in the automated detection and diagnosis involving different types of cyst involving the jaws will be more convincing and beneficial to the dentistry. Many general practitioners still face difficulties in interpreting the radiographs
### Table 1: Characteristics of the included studies

<table>
<thead>
<tr>
<th>S. no</th>
<th>Author (Year)</th>
<th>Application</th>
<th>Radiographic technique</th>
<th>AI tool</th>
<th>Workflow of AI model</th>
<th>Dataset used to develop the AI model</th>
<th>Independent testing dataset</th>
<th>Validation technique</th>
<th>Reference standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Lee (2018)</td>
<td>Dental caries detection</td>
<td>Periapical radiographs</td>
<td>Deep convolutional neural networks</td>
<td>ROI selection (manual); Feature segmentation; detection of lesion; Dental caries diagnosis</td>
<td>Found 1200 caries lesion out of 2400 images in the maxillary premolars/molars</td>
<td>300 caries and 300 non-caries in the form 600 maxillary premolars/molars</td>
<td>Split sample validation</td>
<td>Expert's judgement</td>
</tr>
<tr>
<td>2.</td>
<td>Lee (2018)</td>
<td>Prediction and Identification of teeth with periodontally compromised</td>
<td>Periapical radiographs</td>
<td>Deep convolutional neural network</td>
<td>Image augmentation; Extraction of texture features; Classification of the healthy and periodontally compromised premolars/molars</td>
<td>1392 images exhibiting healthy/periodontally compromised premolars and molars</td>
<td>348 images exhibiting healthy/periodontally compromised premolars and molars</td>
<td>Split sample validation</td>
<td>Clinical and radiological examinations</td>
</tr>
<tr>
<td>3.</td>
<td>Son (2018)</td>
<td>Diagnosis of root fracture, include teeth, decay, missing teeth and periodontal bone resorption</td>
<td>Intraoral and panoramic radiographs</td>
<td>Affinity propagation clustering</td>
<td>Extraction of dental features; Image segmentation; Extraction of features of the segments; Determination of diseases of the segments; Synthesis of the segments; Classification of diseases</td>
<td>87 images of teeth exhibited 16 root fracture, 19 incluse teeth, 17 decay, 16 loss of teeth and 19 periodontal disease with bone resorption</td>
<td>NA</td>
<td>10-fold CV</td>
<td>Expert's judgement</td>
</tr>
<tr>
<td>4.</td>
<td>Hwang (2017)</td>
<td>Osteoporosis detection using structural analysis</td>
<td>Panoramic radiographs</td>
<td>Decision tree; SVM</td>
<td>ROIs selection (manual); Imaging processing; Analysis of texture features; Classification of normal and osteoporotic subjects</td>
<td>454 images from 227 normal and 227 osteoporotic male and female subjects</td>
<td>NA</td>
<td>10-fold CV</td>
<td>DXA examination</td>
</tr>
<tr>
<td>5.</td>
<td>Yilmaz (2017)</td>
<td>Classification of periapical cysts and keratocysts</td>
<td>CBCT</td>
<td>k-NN; Naive Bayes; Decision tree; Random forest; NN; SVM</td>
<td>Lesion detection and segmentation (manual); Extraction of texture features; Classification of periapical cysts and keratocysts</td>
<td>Total of 50 images of subjects 25 tumors and 25 cysts of subjects with cysts/tumors</td>
<td>25 images from subjects with cysts/tumors</td>
<td>Split sample validation</td>
<td>Expert's judgement, radiological and histopathologic examinations 25 images from subjects examinations</td>
</tr>
<tr>
<td>6.</td>
<td>Abdolali (2017)</td>
<td>Classification of dentigerous cysts, keratocysts and radicular cysts</td>
<td>CBCT</td>
<td>SVM; SDA</td>
<td>Lesion segmentation; Extraction of texture features; Classification of lesions</td>
<td>96 images from patients with 38 radicular cysts, 36 dentigerous cysts and 22 keratocysts</td>
<td>NA</td>
<td>3-fold CV</td>
<td>Histopathological examinations</td>
</tr>
<tr>
<td></td>
<td>Author</td>
<td>Title</td>
<td>Methodology</td>
<td>Images/Subjects</td>
<td>Validation</td>
<td>Other Notes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>------------</td>
<td>------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Abdolali (2016)</td>
<td>Maxillofacial cysts segmentation</td>
<td>CBCT Asymmetry analysis Image registration; Asymmetry detection; Cysts segmentations</td>
<td>Total of 97 images involving the subject with 39 radicular cysts, 36 dentigerous cysts and 22 keratocysts</td>
<td>NA</td>
<td>Expert's segmentation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Kavitha (2016)</td>
<td>Osteoporosis based on attributes of the mandibular cortical and trabecular bones</td>
<td>Panoramic radiographs NN Extraction of attributes based on mandibular cortical and trabecular bones; Analysis of the significance of the extracted attributes; Generation of classifier for screening osteoporosis; Classification of normal and osteoporotic subjects</td>
<td>141 images from normal and osteoporotic female subjects aged 45–92 years</td>
<td>5-fold CV</td>
<td>DXA examination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Ngan (2016)</td>
<td>Diagnosis of cracked dental root, incluse teeth, decay, hypodontia and periodontal bone resorption</td>
<td>Dental X-ray images Affinity propagation clustering Extraction of dental features; Image segmentation; Extraction of features of the segments; Determination of diseases of the segments; Synthesis of the segments; Classification of diseases</td>
<td>66 images exhibiting cracked dental root, impacted teeth, decay, hypodontia or periodontal bone resorption</td>
<td>Split sample validation</td>
<td>Expert's judgement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Ohashi (2016)</td>
<td>Detection of the maxillary sinusitis</td>
<td>Panoramic radiographs Asymmetry analysis Edge extraction; Image registration; Detection of maxillary sinusitis; Decision-making (manual)</td>
<td>NA</td>
<td>98 images from 49 subjects with maxillary sinusitis and 49 subjects with healthy sinuses</td>
<td>Independent sample validation</td>
<td>Clinical symptom and CT images</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Rana (2015)</td>
<td>Segmentation and measurement of keratocysts</td>
<td>3D images (MRI/CT) An available navigation software (Brainlab) Identification of keratocysts (manual); Lesion segmentation; Measurement of the lesion volume</td>
<td>NA</td>
<td>38 images from subjects with keratocysts</td>
<td>Independent sample validation</td>
<td>Expert's segmentation and measurement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>Lin (2015)</td>
<td>Identification of alveolar bone loss area</td>
<td>Periapical radiographs Naive Bayes; k-NN; SVM ROI identification (manual); Fusion of texture features; Coarse segmentation of the bone loss area; Fine segmentation of the bone loss area</td>
<td>28 images from subjects with periodontitis</td>
<td>LOOCV/ Split sample validation</td>
<td>Expert's judgement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>Kavitha (2015)</td>
<td>Diagnosis of osteoporosis based on textural features and MCW</td>
<td>Panoramic radiographs Naive Bayes; k-NN; SVM ROI selection (manual); Segmentation cortical margins; Evaluation of eroded cortex; Measurement of MCW; Analysis of textural features; Classification of normal and osteoporotic subjects</td>
<td>141 images from normal and osteoporotic female subjects aged 45–92 years</td>
<td>LOOCV/5-fold CV</td>
<td>DXA examination and expert's measurement</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
with complex diseases, Abdolali et al. proposed an AI model using asymmetry analysis with automatic segmentation of radicular cysts, keratocyst, and dentigerous cysts. Rana et al. effectively made use of available navigation surgical software (Brainlab AG, iPlan, Feldkirchen Germany) for automatic segmentation of keratocysts with measurement of volumes. The last two studies are as follows, Proposed artificial intelligence models that were trained with CBCT images to categorize different cysts and/or tumors (Table 1).

Technically speaking, there are four main steps involved in an AI model’s process of classifying cysts and/or tumors: Detection of lesions, segmentation of lesions, extraction of texture features, and classification of texture features. At the moment, the very first stage of lesion detection in these models still needs to be carried out by hand in order for the models to be able to carry out the subsequent stages of the process automatically. The creation of a fully automated model that is capable of detecting cysts and/or tumors is a difficult task that has not yet been accomplished.

Artificial intelligence for the detection of periodontitis and periapical disease can benefit in identifying alveolar bone resorption and periapical radiolucency, respectively. The alveolar bone loss identification model and the bone loss severity model were both proposed by Lin et al. For the purpose of identifying compromised molars and premolars and predicting hopeless molars and premolars based on the degree of alveolar bone loss occurred, Lee et al. proposed a model with involving deep learning convolutional neural networks. Regarding the diagnosis of periapical disease, Yilmaz et al. projected a model using CBCT images to categorize periapical cysts.

Many studies have developed models to detect caries using nonclinical 2D images from extracted teeth, but only one study in this review used clinical X-ray images. Although the models’ diagnostic performance was satisfactory in these preclinical studies, the results may be overly positive because the training-testing images featured noticeable lesions from extracted teeth but lacked images of other oral tissues. Maxillary premolars and molars are particularly vulnerable to caries, so Lee et al. proposed a model for caries detection that uses deep learning algorithms.

The other studies included in this area are AI models to aid in the detection of maxillary sinusitis and multiple dental diseases. This is further evidence that various areas of dental and maxillofacial radiology are beginning to investigate the use of artificial intelligence techniques. The findings of the current systematic review of diagnostic accuracy studies were remarkably heterogeneous, attributed to differences in the conceptualization and execution of included studies. Hence a meta-analysis was not feasible. As a result, careful evaluation of the quality of included studies was warranted. The Agency for Healthcare Research and Quality, the Cochrane Collaboration, and the National Institute for Health and Clinical Excellence endorsed the use of the QUADAS tool in the systematic reviews of diagnostic accuracy studies. In the present review, the studies on cadavers and extracted specimens were deemed to be at high risk for patient selection and index test under the applicability domain which was in congruence with a recent systematic review.

Future studies may further explore other potential options to integrate additional information by other imaging technologies such as MRI and image fusion methods in significant asymmetric lesions of cysts and tumors without distinctive boundaries. Digital radiography could also be used to increase the accuracy of textural parameters in osteoporosis screening. CNN techniques based on deep learning with high-resolution large-scale images would show good precision and discriminatory potential. These systems add tremendous value to the clinical practice by strengthening diagnostic accuracy, optimizing clinical management, and determining treatment prognosis. Thereby, can assist clinicians in delivering the highest possible treatment.

This current systematic review focused solely on the role of AI in dental and maxillofacial radiology in diagnosing a wide range of oral and maxillofacial diseases, so it excluded studies that used standard oral radiology for AI claims but then again had no direct significance for specific oral diseases. The articles identifying cephalometric landmarks were specifically left out of the present review. Because of this, we were able to devote more time to researching how to use AI to identify oral health issues. While the described artificial intelligence models show promise, more work is needed to ensure their generalizability and reliability using larger institutional data. Also, deep learning, which is considered to understand the most advancement in AI technique, deep learning model should be used to make diagnostic AI models in the field of medicine and should be implemented in dental and maxillofacial imaging. In the field of dental and maxillofacial radiology, one of the long-term goals of AI research and development is to not
only bring the performance of AI models up to the level of that of specialists but also in identifying or detecting very early lesions which cannot be observed with human eyes.

**Conclusion**

In conclusion, AI models described in the included studies showed a wide range of potential applications for the diagnosis of common oral diseases, including osteoporosis, the classification/segmentation of maxillofacial cysts and/or tumors, the detection of dental caries, multiple dental diseases, maxillary sinusitis, and periodontitis/periapical disease. Therefore, implementing AI for patient diagnosis and clinical decision-making is feasible, and the technology should be seen as a trustworthy modality for future applications in oral diagnosis that can aid dental practitioners in a variety of ways, including cutting down on chair side time, saving extra steps in routine diagnosis protocol, achieving excellent infection, and delivering quality treatment with accuracy and precision. For future systematic reviews to accurately describe and evaluate the value and impact of AI in daily practice, a greater emphasis on these specific areas of oral diagnosis is required. Until adequate, multi-institutional images are used to validate the generalizability and reliability of these models, the diagnostic performance of the AI models will continue to vary depending on the algorithms used.

**References**


28. Devito KL, de Souza Barbosa F, Felippe Filho WN. An artificial multilayer perceptron neural network for diagnosis of proximal...